SANDIA REPORT

SAND2006-6890
Unlimited Release
October 2006

Updating Time-To-Failure Distributions Based On Field Observations and Sensor Data

Daniel Briand, Kelly S. Lowder, and Donald N. Shirah

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@adonis.osti.gov
Online ordering:	http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.fedworld.gov
Online order:	http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2006-6890
Unlimited Release
October 2006

Updating Time-To-Failure Distributions Based On Field Observations and Sensor Data

Daniel Briand, Kelly S. Lowder, and Donald N. Shirah
System Sustainment and Readiness Technologies
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS1101

Abstract

Enterprise level logistics and prognostics and health management (PHM) modeling efforts use reliability focused failure distributions to characterize the probability of failure over the lifetime of a component. This research characterized the Sandia National Laboratories' developed combined lifecycle (CMBL) distribution and explored methods for updating this distribution as systems age and new failure data becomes available. The initial results obtained in applying a Bayesian sequential updating methodology to the CMBL distribution shows promise. This research also resulted in the development of a closed-form full life cycle (CFLC) distribution similar to the CMBL distribution but with slightly different, yet commonly recognized, input parameters. Further research is warranted to provide additional theoretical validation of the distributions, complete the updating methods for the CMBL distribution, evaluate a Bayesian updating methodology for the CFLC distribution, and determine which updating methods would be most appropriate for enterprise level logistics and PHM modeling.

CONTENTS

FIGURES

TABLES

NOMENCLATURE

CMBL	Combined Lifecycle
DOE	Department of Energy
SNL	Sandia National Laboratories
PHM	Prognostics & Health Management
TTF	Time to Failure
LDRD	Laboratory Directed Research and Development
SoSAT	System of Systems Analysis Toolset
RTCE	Real Time Consequence Engine
SEM	Support Enterprise Model
SoS	System of Systems
FCS	Future Combat System
SMO	State Model Object
MTTF	Mean Time To Failure
MTTR	Mean Time To Repair
CFLC	Closed-form Full Life Cycle

1. INTRODUCTION

One of the primary purposes of current enterprise level modeling efforts is to use component/system reliability estimates along with inventory levels, maintenance and inspection schedules, and operational requirements to optimize supply/repair chain processes. In addition, prognostic and health management (PHM) modeling uses component/system reliability estimates as a baseline from which the data from sensors and maintenance events along with data fusion techniques can determine component health trends. These component health trends may help predict failure far enough in advance to be able to modify operations and maintenance schedules for the purposes of maximizing system availability or minimizing maintenance and spares costs. In either case, once a component's lifecycle in terms of failure probability is characterized, a methodology for how to update that characterization based on the availability of new data is required.

Reliability models depend on failure distributions to characterize the probability of failure over the lifetime of a component. Many types of components typically will have a bathtub-shaped failure rate life distribution. This distribution is commonly characterized by a decreasing failure rate during the early portion of a component's life, a constant failure rate during the useful portion of its life, and an increasing failure rate during the wear-out portion of its life, as shown in Figure 1. During the early portion of its life, failures are typically caused by manufacturing defects. During the useful portion of its life, component failures are usually caused by chance, perhaps as a result of overstress or a shock to the system. The wear-out portion of its life is characterized by wear or accumulated damage that exceeds allowable limits for normal operation [1]. In many supply/repair models, the failure rate distributions for components model only the useful life period, typically with a constant failure rate that does not take into account the aging process and the wear out problems that will occur [2]. However, when modeling at the unit or enterprise level, being able to use the failure characteristics across a component's lifetime provides greater accuracy and usefulness of logistics models.

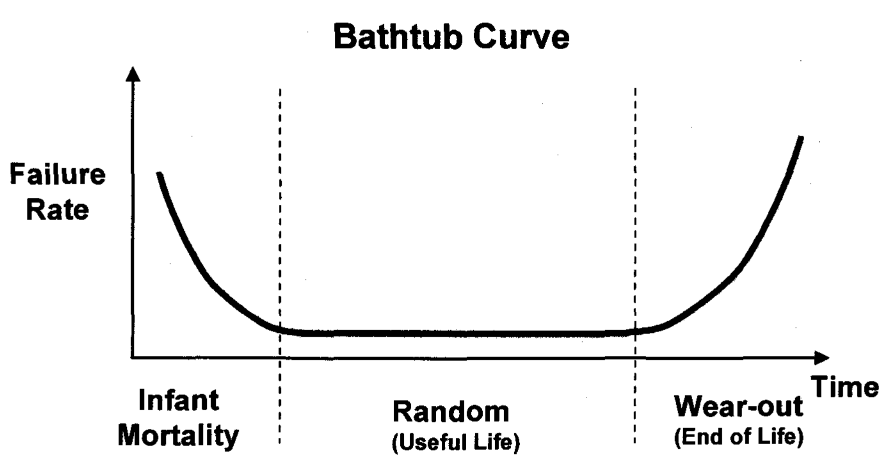


Figure 1. The Bathtub Curve [3].

Knowing how the failure rate will change over time is also important to being able to predict a failure of an individual component early enough to be able to modify operations and maintenance scheduled in order to maximize availability. This capability is commonly termed

PHM. The prognostic capability relies on knowing how an individual component's failure rate deviates from the "average" or expected component's failure rate distribution. The change can be analyzed to determine or predict the component's remaining useful life. This prognostic capability relies on sensors operating in real-time and/or inspections to detect changes in a component's health, and data fusion algorithms that use that information to predict the change in time-to-failure (TTF) or remaining useful life. Whether to optimize the supply and repair chain process or to implement an effective PHM program, the accurate portrayal of a component's failure distribution across its entire lifecycle is critical to maximizing a system's availability while minimizing parts and maintenance costs.

The goal of this Laboratory Directed Research and Development (LDRD) effort was to better understand how to correctly update TTF distributions, based initially on sparse data, data from similar components, and expert opinion, with new observations and sensor data. The remainder of this paper will discuss updating bathtub shaped TTF distributions. More specifically, Section 2 will describe the Sandia developed combined lifecycle (CMBL) distribution used for enterprise level logistics and PHM component reliability representation. Section 3 will describe the Bayesian sequential updating approach that treated each section of the CMBL distribution separately. Section 4 presents the development of a closed form solution to a slight variation of the CMBL distribution. Finally, Section 5 summarizes the results and suggests areas for further exploration.

2. COMBINED LIFECYCLE (CMBL) DISTRIBUTION

2.1. Background

Early in a system's life cycle, reliability data may not be abundant, as opposed to later in the system's lifecycle, where operational failure data becomes available. Where data is not abundant, expert opinion is solicited. Sometimes data from similar components can be used but must be updated with expert opinion. The expert opinion may come from engineers/technicians who are typically not statisticians or reliability experts so it helps immensely to be able to elicit the necessary information using more common terms and concepts, i.e., how long is the burn-in phase, what percent of total component failures are a result of burn-in, what is the mean life expectancy of the component given it makes it to its wear-out phase, etc. Trying to use common failure distributions, such as combinations of the Gamma and/or Weibull distributions, may be quite involved since converting expert opinion to a parameter value would most likely require several iterative steps [4]. Despite considerable published works in bathtub shaped failure distributions, few practical models are available [5].

2.2. How It Is Used Today

The CMBL distribution, developed by Dr. James E. Campbell, Sandia National Laboratories and Dr. Dennis Longsine, Intera Corporation, is used in two key simulations, the System of Systems Analysis Toolset (SoSAT) and the Real Time Consequence Engine (RTCE). Both simulations are in continuing development at Sandia National Laboratories and use the CMBL distribution as one of six different distributions that can be used to model component TTF. Other logistics simulations such as the Support Enterprise Model (SEM), a multi-echelon supply and repair chain optimization model currently being used to support multimillion dollar business case decisions for the Joint Strike Fighter, may use the CMBL distribution in the future. Hence, there is a need to evaluate the CMBL distribution's mathematical properties and methodologies to update the distribution as new data becomes available.

2.2.1. System of Systems Analysis Toolset (SoSAT)

The primary focus of SoSAT is to support system of systems (SoS) analyses for the U. S. Army's Future Combat System (FCS) sustainability requirements, which include maximizing available warfighting capabilities while reducing logistics footprint and maintenance personnel [6]. SoSAT has a multi-system time simulation capability that uses State Model Objects (SMO) to enable a system, its elements, and its functionality to be encapsulated for use in the simulation (Figure 2). Every system in the simulation is represented by a SMO which models the system's functionality. Controlling simulation software provides needed information on environmental conditions, terrain, use conditions, supply network information, etc. A scenario model describes the detailed scenarios that the systems will follow during the simulation. A combat damage model provides a mechanism to simulate the effects of combat damage to the individual system primary elements or damage that completely disables the system. A supplies and services model provides a means for spare parts and consumables to move from system to system in the simulation and makes maintenance services available to systems requiring repairs [7].

Figure 2. Multi-system Simulation Concept [7].

2.2.2. Real Time Consequence Engine

The purpose of the Real Time Consequence Engine (RTCE) is to conduct consequence analysis based on updated component failure information made available by prognostic data fusion schemes. More specifically, the RTCE will predict in real-time the effect that operational strategies, as well as repair/replace/inspect/wait strategies, will have on the modeled system if that action is taken at the time changes in the reliability of the system are detected. This consequence analysis part of the prognostic problem must be performed in such a way that the expected benefit from performing each alternative operational/maintenance strategy can be calculated. Then, given a set of alternatives each with an expected cost/benefit, the alternative (or subset of alternatives) can be chosen that provides the optimal benefit, such as maximizing system availability while minimizing cost.

The RTCE is a forward looking simulation that can help analyze the consequences of various operational and maintenance strategies once a change in a component's remaining useful life or time-to-failure has been detected (Figure 3). The RTCE takes the system health predictions from prognostic analysis methods and develops operational projections into the future, i.e., what will be the overall impact on the system if a certain impending failure mitigation action is taken. Simulation is used to model the long-term impact of possible maintenance/inspection strategies for each failure mode on overall system performance and cost.

The simulation mimics the reliability behavior of equipment in terms of simulated equipment failures, repairs, scheduled maintenance, and inspections. The simulation is based on user-definable maintenance and inspection schedules and a reliability model with time-to-failure and time-to-repair distributions for all failure modes. A spares model is included since the availability of spares may be a major factor in decision-making when a pending failure is identified. By running repeated simulations, the RTCE can be used to analyze possible scenarios accounting for projected TTF, parts availability, planned equipment use schedule, and mission profiles. Performance metrics such as mean time between failures (MTBF), mean time to repair (MTTR), availability, maintenance cost, downtime cost, etc., are calculated. Thus, by running the simulation with different operational settings and maintenance schedules, it can be used to examine the consequences of alternative equipment use and maintenance scenarios [8].

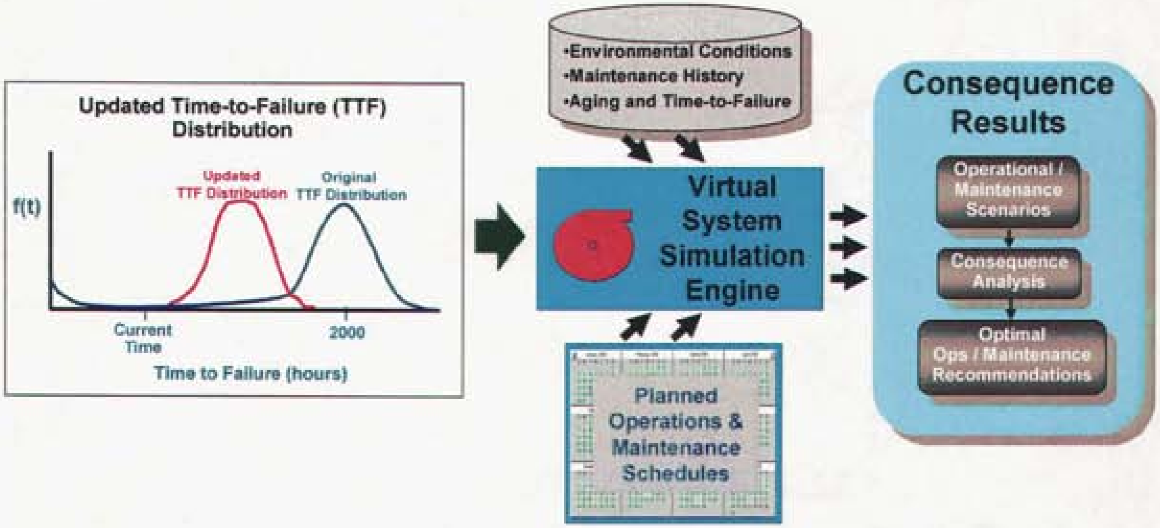


Figure 3. Real-Time Consequence Engine.

2.3. Mathematical Formulation

The CMBL distribution is a sectional model that assumes a linearly declining failure rate during infant mortality, a constant failure rate during the random failure section or useful life, and a normally distributed TTF during wear-out in the end of life phase (Figure 4). The distribution is characterized by five parameters that are:

1. The mean of the normally distributed portion of the TTF distribution.
2. The standard deviation of the normally distributed portion of the TTF distribution.
3. The probability that the component will fail during burn-in.
4. The duration of the burn-in portion of the distribution.
5. The probability that failure will occur randomly after burn-in.

This distribution as formulated must include the random and normal wear-out TTF portions, but can exclude the infant mortality portion [9].

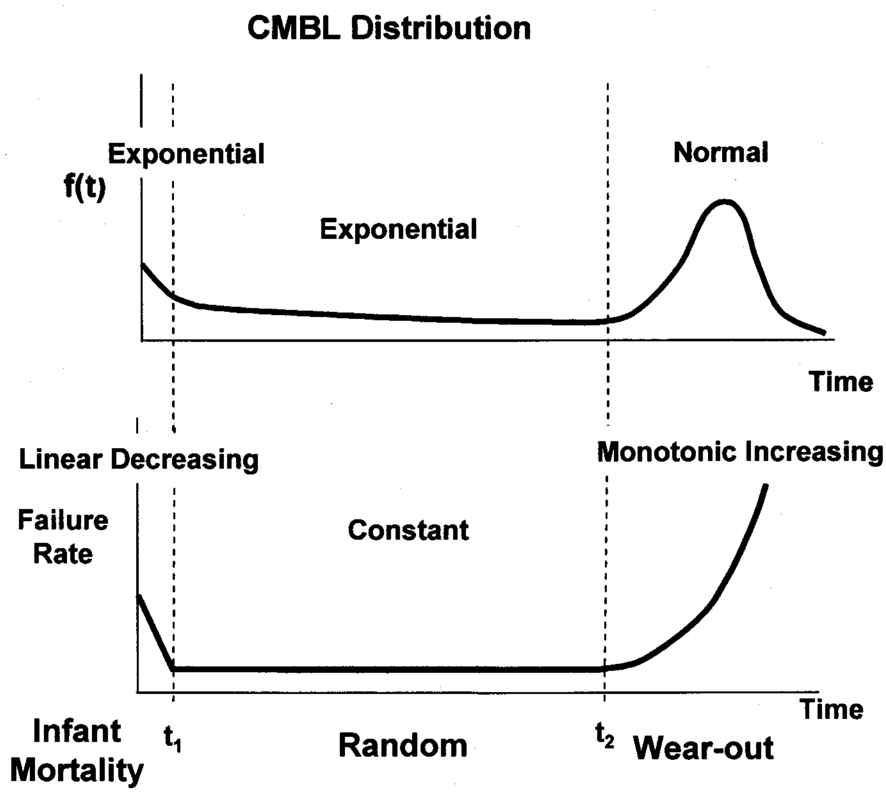


Figure 4. CMBL PDF and Time Dependent Failure Rate Distribution

The explicit form of the distribution is:

$$
f(x) = \begin{array}{lll} \lambda_d e^{-\lambda_d t} & 0 \le t \le t_1 & \lambda_d = (mt + b \\ \lambda_c e^{-\lambda_c t} & t_1 \le t \le t_2 & \\ \dfrac{A}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(t-u)^2} & t_2 \le t \le \infty & \end{array} \qquad (1)
$$

where

μ = mean of the normally distributed portion of the TTF distribution.
σ = standard deviation of the normally distributed portion of the TTF distribution.
λ_d= The failure rate for the linearly decreasing failure rate portion
λ_c = The failure rate for the constant failure rate portion.
t_1= burn-in duration.
t_2 = transition from constant failure rate to the normal TTF portion.
A = area multiplier for the truncated normal portion.
F_1 = fraction of the distribution area for the infant mortality portion.
F_2 = fraction of the distribution area for the constant failure rate portion.

The unknowns in the above parameters are A, λ_c, λ_d, and t_2. Since this is a pdf, the total area under the distribution must be 1 such that:

$$1-F(t_2)=1-F_1-F_2 \quad \Rightarrow \quad F(t_2)=F_1+F_2 \tag{2}$$

However, since λ_c and t_2 are not known, $F(t_2)$ is also not known. A factor A is applied to allow a difference between each side of Equation (1) such that:

$$A(1-F(t_2))=1-F_1-F_2 \quad \Rightarrow \quad A=\frac{1-F_1-F_2}{1-F(t_2)} \tag{3}$$

where $F(t_2)$ is the area under the normal from $-\infty$ to t_2. The area under the constant failure rate portion of the TTF is F_2 such that:

$$F_2=\int_{t_1}^{t_2}\lambda_c e^{-\lambda_c t}dt \quad = e^{-\lambda_c t_1}-e^{-\lambda_c t_2} \tag{4}$$

Solving Equation (3) for t_2 results in

$$t_2=\frac{-\ln\left(e^{-\lambda_c t_1}-F_2\right)}{\lambda_c} \tag{5}$$

Ensuring that the TTF be continuous as it transitions from the constant failure rate portion to the normally distributed portion requires:

$$\frac{A}{\sigma\sqrt{2\pi}}e^{\frac{-(T_2-\mu)^2}{2\sigma^2}}=\lambda_c e^{-\lambda_c T_2} \tag{6}$$

An iteration scheme is set up that increments the λ_c and adjusts A and t_2 appropriately. The scheme uses the current value of λ_c in Equation (5) to get t_2, then t_2 is substituted into Equation (3) to get A. The new values of A, t_2, and λ_c are inserted into Equation (6) and this process continues until Equation (6) is essentially true while ensuring that $f(t)$ integrates to 1.

The infant mortality portion of the PDF is characterized by a linearly decreasing failure rate. In general, the failure rate from time 0 to t is:

$$\lambda_d(t)=mt+b \tag{7}$$

and modeling the infant mortality portion of the CMBL distribution as an exponential distribution implies:

$$f(t_1)=\lambda_d e^{-\lambda_d t} \quad or \quad f(t_1)=(mt+b)e^{-(mt+b)t} \tag{8}$$

Ensuring that the TTF is continuous as it transitions from the linearly decreasing failure rate portion to the constant failure rate portion requires:

$$\lambda_d(t_1)=mt_1+b=\lambda_c \tag{9}$$

Also,

$$F(t_1)=1-e^{-\int_0^{T_1}(mt+b)dt}=1-e^{-\left(\frac{mt_1^2}{2}+bt_1\right)} \tag{10}$$

Equation (9) is used to solve for b and substitute into (10) to get m. At $t = t_1$, the linearly declining failure rate must equal the constant failure rate λ_c.

$$b = \lambda_c - mt_1 \tag{11}$$

Also at t_1, the failure probability must be $F(t_1)$. From Equation (11), this gives

$$F(t_1) = 1 - e^{-\left(\frac{mt_1^2}{2} + (\lambda_c - mt_1)t_1\right)} \tag{12}$$

and

$$\ln(1 - f_1) = \frac{mt_1^2}{2} - \lambda_c t \tag{13}$$

and

$$m = \frac{2[\ln(1 - f_1) + \lambda_c t_1]}{t_1^2} \tag{14}$$

Once the CMBL distribution is characterized using scarce data, data from similar components, and expert opinion, the next step is to update it as new data from the components becomes available. As the system being modeled undergoes extensive testing and is used under normal operations, data on the number of failures, mean time to failure (MTTF), mean time between failures (MTBF), mean time to repair (MTTR), etc., becomes available. It is appropriate throughout the modeling process to update the parameters of the original CMBL distribution in some fashion to improve its accuracy. Updating the distribution should occur throughout a component's lifecycle since improvements in the component and changes in its use may alter its inherent reliability.

In addition, the resulting updated failure distribution must be in a usable format for the supply/repair chain or PHM model for which it resides since the resulting distribution will be fed back into the supply/repair chain or PHM model. An empirical distribution is possible, but it would be better to have the updated distribution be the same as the originating distribution. However, in some models it may not make much difference as long as the resulting distribution can be easily updated as new data becomes available. Being able to transition the CMBL distribution to a more common form such as a single commonly used distribution would most likely not work. Transition from the CMBL distribution to a bathtub distribution modeled by two or more distributions such as three Weibull distributions (one for each section) would present even greater challenges. One approach to updating the CMBL distribution with the intent to obtain the parameters of the original distribution is presented in Section 3.

To help in the evaluation of the CMBL distribution, a CMBL-Distribution Grapher written in Visual Basic.NET (VB.NET) model was created. This model provided the capability to view the resultant distribution based on the user provided five input parameters as described above. The CMBL-Distribution Grapher also provided data streams from the distribution for checking convergence when evaluating the Bayesian updating by section method. Finally, the CMBL-Distribution Grapher provided an automated way to evaluate the limits of the distribution. Additional details of the CMBL-Distribution Grapher can be found in Appendix A.

3. BAYESIAN UPDATING BY SECTION APPROACH

3.1. Introduction

The approach to updating the CMBL distribution, described in the following paragraphs, assumed that each section of this piecewise continuous distribution, infant mortality, random, and wear-out sections, could be updated independently of the other. This approach assumed an appropriate conjugate prior depending upon where the new data occurred, i.e., 0 to t_1, t_1 to t_2, and t_2 to infinity, where the new data has the same underlying distribution as the section it occurred in, and used a Bayesian updating methodology to determine the posterior distribution. This posterior distribution, or an estimate of the posterior distribution, was used as the subsequent prior as new failures occur in each section. For example, a Gamma (α,β) was used as the prior distribution in the random (constant failure rate) section of the CMBL distribution and the new data was assumed to follow an Exponential distribution with a constant failure rate. To determine the resulting posterior distribution, a modified iterative scheme determined the new constant failure rate λ_c and subsequently the new parameters of the CMBL distribution, which are then used as the next prior distribution in the sequential updating scheme.

3.2. Approach

This method determined if the CMBL distribution could converge to the distribution of the data through a section by section Bayesian updating methodology. As mentioned earlier, each new data point was assumed to have the same underlying failure distribution as the section of the CMBL distribution where it occurred. Several additional simplifying assumptions were also made. Only the random and wear-out portions of the bathtub curve were evaluated specifically, although the infant mortality portion was evaluated briefly in the process to determine if there would be any unusual behavior in the transition region around t_1. It was assumed that with a Bayesian updating methodology, the CMBL distribution would be updated one data point at a time, thus the reference to sequential updating. Since the new data created from different input parameters (or actual failure occurrences in the field) can be quite extreme, especially with regard to the Exponential distribution, a weighting scheme was introduced to "smooth" the convergence process. Finally, the standard deviation for the data distribution in the wear-out portion of the distribution was assumed to be known and held constant. The validation of this approach is broken down into several steps as outlined in the following paragraphs.

3.2.1. Step 1

The first step was to evaluate the process for updating the CMBL distribution with data falling within the wear-out section of the distribution. This was a relatively straightforward process. In a Bayesian updating methodology, the conjugate prior is a Normal (u, σ). Using a Normal $(u_{CMBL}, \sigma_{CMBL})$ for the prior and a Normal (u_0, σ) for the single new data point results in a Normal (u_1, σ_1) where

$$u_1 = \frac{\frac{u_{CMBL}}{\sigma_{CMBL}^{2}} + \frac{t_1}{\sigma^2}}{\frac{1}{\sigma_{CMBL}^{2}} + \frac{1}{\sigma^2}} \quad (15)$$

This posterior Normal (u_1, σ_1) distribution becomes the next prior distribution, and this updating process is repeated for each new data point. As expected, this process results in the Normal (u_{CMBL}, σ_{CMBL}) converging to the Normal (u_d, σ) reasonably well.

3.2.2. Step 2

The second step was to evaluate the Bayesian updating process for the data points that fall within the random failure section of the bathtub distribution. The only user input available for this section is F_2, the fraction of failures occurring in the random failure portion, but the resulting λ_c, the failure rate for the constant failure rate portion, is determined within the iterative process of the CMBL distribution. Since the data occurring in the random section is being modeled by an Exponential (λ) where λ is considered a random variable, a typical conjugate prior used in many reliability applications is the Gamma (α, β). Since the initial CMBL estimate and subsequent priors are being taken as a single instance of failure, the Bayesian prior is the Gamma (1, β), which is essentially an Exponential (β) as given by:

$$g(\lambda \mid \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \lambda^{\alpha-1} e^{-\lambda/\beta} \quad (16)$$

$$where\; \alpha = 1 \quad => g(\lambda : 1, \beta) = \frac{1}{\beta} e^{-\lambda/\beta} \quad (17)$$

so $\lambda \sim Expon(\beta)$. This results in a posterior distribution that is:

$$g(\lambda \mid t, \alpha, \beta) = \frac{\frac{1}{\Gamma(\alpha)\beta^{\alpha}} \lambda^{\alpha-1} e^{-\lambda/\beta} \lambda e^{-\lambda t}}{\int \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \lambda^{\alpha-1} e^{-\lambda/\beta} \lambda e^{-\lambda t}} \quad (18)$$

which becomes:

$$g(\lambda \mid t, \alpha, \beta) = \frac{1}{\Gamma(\alpha+1)(\frac{\beta}{\beta t+1})^{\alpha+1}} \lambda^{\alpha+1-1} e^{-\lambda/\beta} \lambda e^{-\lambda(t+1/\beta)}, \quad 0 < \lambda < \infty \quad (19)$$

with an expected value of:

$$E(\lambda \mid t, \alpha, \beta) = \frac{\beta(\alpha+1)}{\beta t+1} [10] \quad (20)$$

In this case, assuming successive single data point Bayesian updating, the result is a Gamma (2, $\beta_0/\beta_0 t_1+1$). The resulting $E_{Gamma}(\lambda \mid t, 2, \beta)$ is the new estimate for λ_c, which is used in the CMBL iterative process to calculate a new F_2. This change in F_2 is then used with the other unchanged parameters to the CMBL distribution, which now becomes the prior distribution for the next data update. Using this estimate alone did not provide accurate results when using the successive single data point updating methodology. This shortfall became apparent when the data represented the original distribution and the sequential updating process resulted in a mean that was not reasonably close to the expected mean. However, when Equation (20) was multiplied by a factor of approximately 0.63 and data represented the original distribution derived from the input parameters, the deviation from the expected mean essentially disappeared. Further evaluation of the 0.63 factor is presented in a subsequent section.

3.2.3. Step 3

The third step examined updating both the random and wear-out section of the CMBL distribution simultaneously. The primary goal was to determine if the resulting Bayesian updating scheme for both sections simultaneously resulted in a convergence to the distribution of the data. Particular attention was paid to the transition region around t_2 since this requires an iterative scheme to ensure the failure distribution of the two sections remain essentially continuous.

3.3. Wearout Section Updating

Several evaluation runs (using only the random and wear-out portions of the CMBL distribution) were made to determine the characteristics of convergence to the distribution of the data. As an example, the Baseline ($\mu = 200$), $\mu = 250$, and $\mu = 150$ CMBL distribution sets of parameters are shown in Table 1. The baseline distribution provides the starting or initial prior distribution. Using data from the $\mu = 250$ set of parameters, which changes μ only, focuses the convergence on updating the normal portion of the CMBL distribution although the random section has to adjust as well. Using data generated from the $\mu = 150$ set of parameters, which again changes μ only, focuses the convergence on updating the random portion of the CMBL distribution although the normal section has to adjust appropriately.

Table 1. Baseline, μ = 250, and μ = 150 Input Parameters

	Baseline	**μ = 250**	**μ = 150**
μ	200	250	150
σ	20	20	20
F_1	0	0	0
F_2	.2	.2	.2
t_1	0	0	0

3.3.1. Baseline Input Parameters

To give an initial validation of the updating approach, the Baseline input parameters for the CMBL distribution were used as prior information and sequentially updated with data that represented the same Baseline distribution. The results of each update were used as the prior information for the next update. As mentioned earlier, a weighting scheme, that essentially provided a user defined fraction of the change between the old mean and the new data point, was used to moderate extreme new data values. Convergence to the distribution of the data generated from the Baseline input parameters was consistent for the entire 1000 iterations as shown in Figure 5. In converging from a mean of 200 to the data mean of 200 (actually the mean converged to 200.4 if the last 500 iterations are averaged), the t_2 also converged to the value of 155, very close to the expected value of 154 (based on an average of the last 500 iterations. A moving average of the data, averaged over six data points, shows the variability in the data representing the two sections of the bathtub distribution, random and wear-out.

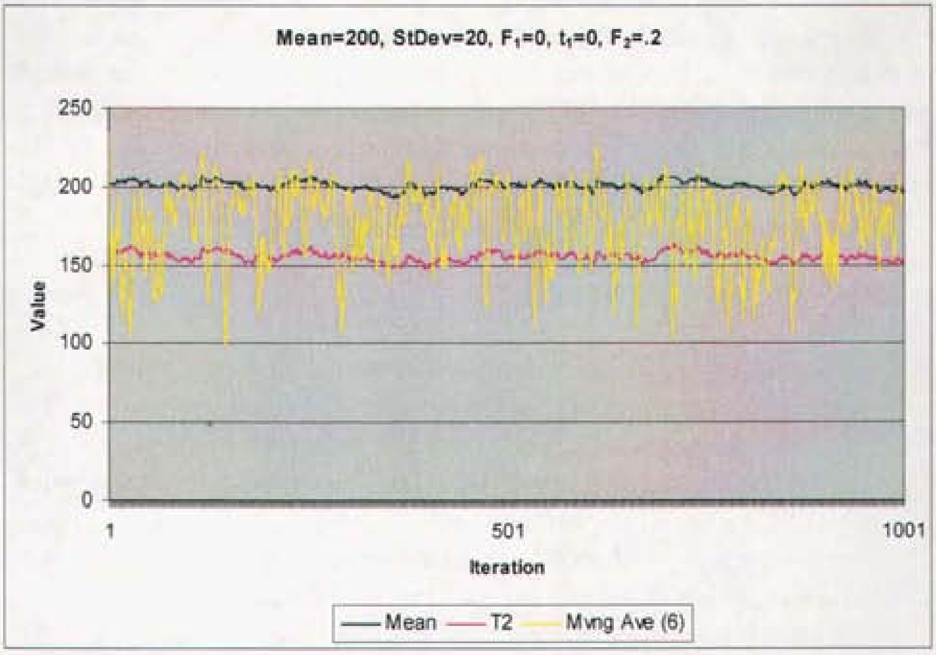


Figure 5. Convergence to Baseline Data

3.3.2. μ = 250 Input Parameters

Starting with the baseline input parameters for the CMBL distribution, successive CMBL distributions with their updated parameters were created using the new data (failure times) from a CMBL distribution with the $\mu = 250$ input parameters. Convergence to the distribution of the data generated from the $\mu = 250$ input parameters was relatively consistent and quick. In converging from a mean of 200 to a mean of 250, the t_2 also converged to the anticipated value of 202, from the starting value of 154. Convergence appears to occur within about 200 iterations, as shown in Figure 6. A moving average of the data, averaged over six data points, shows the variability in the data representing the two sections of the bathtub distribution, random and wear-out.

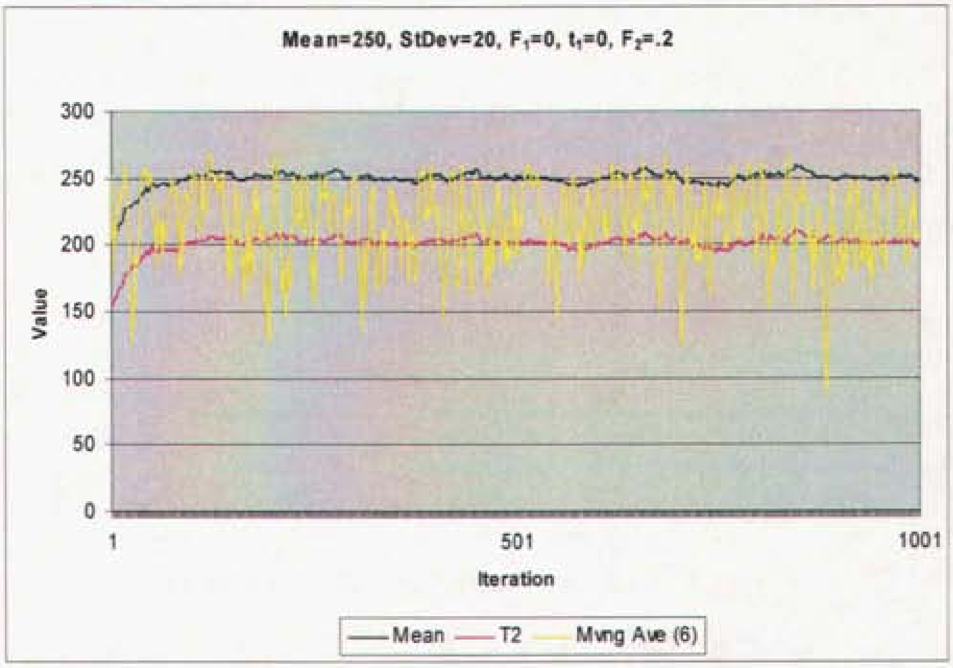


Figure 6. Convergence to μ = 250 CMBL Data

3.3.3. μ = 150 Input Parameters

Convergence to the distribution of the data generated from the μ = 150 input parameters was relatively consistent but not as quick as with the μ = 250 input parameters. It took about 250 iterations to reach near the new mean of 150, as shown in Figure 7. The t_2 converged to the anticipated value of 107. Again, a moving average of the data, averaged over six data points, shows the variability in the data representing the two sections of the bathtub distribution, random and wear-out.

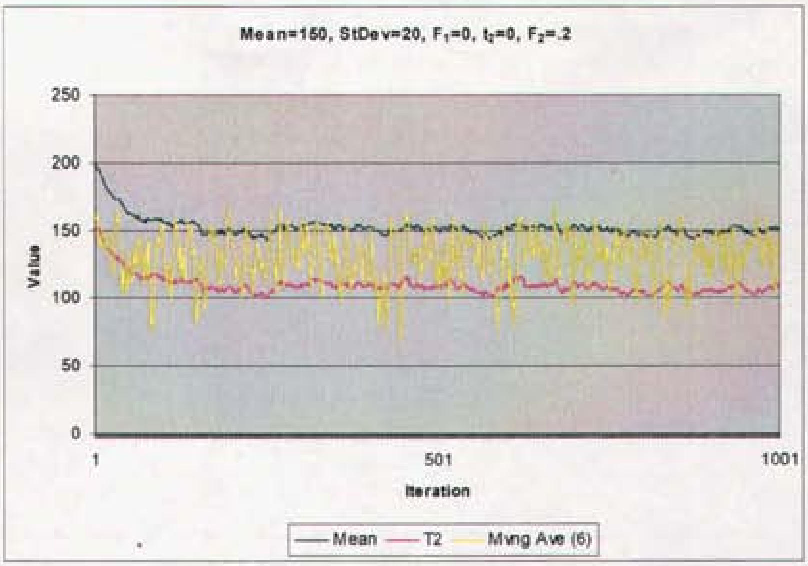


Figure 7. Convergence to μ = 150 CMBL Data

A comparison of the results of the Baseline, $\mu = 250$, and $\mu = 150$ parameters is shown in Figure 8. The comparison shows an appropriate shift in the mean of the distribution while an appropriate shift in λ_c occurs in response to the shift in the mean. For example, when the mean of the distribution of the data causes a shift from 200 to 150, the failure rate λ_c of the random section of the CMBL distribution increases to ensure F_2 remains at constant at 0.2. The opposite occurs when the mean of the distribution of the data shifts from 200 to 250, as expected. It is important to note that this approach works in this application only because of the iteration scheme calculations that increment the λ_c and adjusts A and t_2 appropriately to both ensure virtual continuity and guarantee that the probability distribution integrates to one.

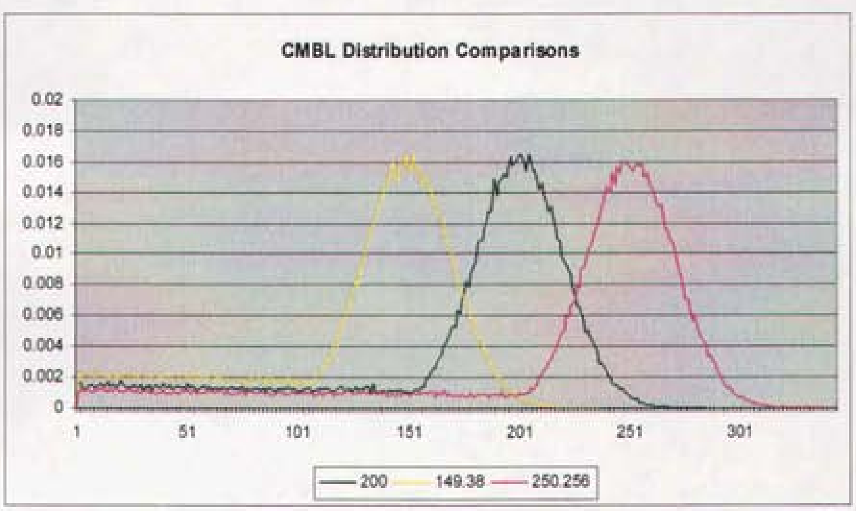


Figure 8. Baseline, μ = 250, and μ = 150 Comparisons

3.3. Random Section Updating

The next step in the evaluation of this approach was to hold the mean constant and vary F_2, the fraction of failures occurring in the random failure section. This step also helped determine the underlying need for the .63 factor discussed earlier. In this step, attempts were made to update the Baseline CMBL distribution with data that had a change only in F_2, the fraction of failures due to random failure. Unfortunately, this step was never completed successfully, either the updating scheme results caused the calculation of the CMBL distribution to "blow up" or the posterior distribution never converged to the distribution of the data, where either F_2=.1, or F_2=.3 as shown in Table 2.

Table 2. Baseline, F_2=.1, and F_2=.3 Input Parameters

	Baseline	F_2=.1	F_2=.3
μ	200	200	200
σ	20	20	20
F_1	0	0	0
F_2	.2	.1	.3
t_1	0	0	0

The primary reason for the problem in this area stemmed from the fact that the data being updated in the random section of the CMBL distribution was in effect being applied to a truncated Exponential distribution with limits $0 < t \leq t_2$ versus a full Exponential distribution with limits $0 < t < \infty$. More specifically, the λ_c = .00145 in this example represented a MTBF of about 690 with respect to the full Exponential distribution. Since the mean of the data for this example was 200, with a standard deviation of 20, the mean of the random portion was well outside virtually all occurrences of failure. In addition, the updating scheme allowed only the data occurring within the random portion of the distribution (any data values $< t_2$) to be updated using the random updating process described earlier. Any data occurring in the wear-out section of the distribution would be greater in magnitude than that in the random section. As a result, only the smaller TTFs would be updated in the random section. Since the updating scheme in the random section used Equation (20), which is essentially an average of the previous failure rate and the inverse of the new data, the resultant λ_c would relatively quickly be driven so high that the CMBL distribution calculation limitations would be exceeded, i.e., F_2 would become too large. The approximate 0.63 factor in this example served to keep the λ_c in a range that allowed the updating procedure to run through to convergence and somewhat beyond.

Several unsuccessful approaches to remedy this problem were attempted. First, the data occurring in the random section of the CMBL distribution during the updating process was transformed from the truncated Exponential distribution with the new data value and the old t_2, back to a full Exponential distribution with a range of $0 < t < \infty$. The resultant transformed new data point then was more appropriately exponentially distributed with a λ_c= 690 (in this example) as opposed to being less than t_2= ~ 154. However, the probability distribution of the random section is relatively flat and as a result, the resultant transformed data was skewed to the right as one would expect, but this still caused the resultant CMBL distribution calculation limitations to be exceeded.

Realizing that the probability distribution in the random section was relatively flat, a second attempt took the proportion of the new data point t to the range from $0 < t \leq t_2$, and applied that proportion to the old λ_c with a range $0 < t \leq \infty$. More specifically, this approach assumed the data falling into the random section of the distribution was uniformly distributed over the range $0 < t \leq t_2$, and based on the mean of that interval, transformed the data to a uniformly distributed interval with a mean equal to the inverse of λ_c. This transformation appeared to work better than previous attempts although a measure of skewness still occurred. Typically a larger data set could be evaluated before the limits of the CMBL distribution calculations were exceeded. Even with the larger amount of data that could be updated, convergence to the distribution of the data could not be confirmed.

3.4. Summary

Additional work is needed on this approach that uses a Bayesian updating methodology while treating each section of the CMBL distribution separately. The iterative process inherent in the CMBL distribution calculations allow updating of the wear-out section of the distribution while keeping F_2 constant. In this case, convergence to the distribution of the data was shown to be quick; quick enough for use in supporting enterprise level modeling where field data is used to update the original CMBL parameter estimates.

However, the results of trying to update the random section of the CMBL distribution while keeping the mean of the wear-out section constant were not as successful. The transformation process which assumes the data occurring in the random section of the distribution is uniformly distributed showed promise but additional research is needed. An extension of this approach may be to reverse this transformation process, determine the truncated λ_c, update the truncated λ_c with the data, then once again reverse the transformation back to a full Exponential distribution with a range of $0 < t < \infty$, then determine F_2 in preparation for the next updating event. In this case, determining the input parameters may require considerable reverse engineering of the iteration procedure.

4. CLOSED-FORM FULL LIFE-CYCLE SOLUTION

4.1. Background & Motivation

Upon review of the piecewise-defined CMBL distribution, it was immediately apparent that this distribution had the right qualities to model the probability of failure throughout the three phases of a part's life cycle, namely, infant mortality, random failure, and wear-out. However, the piecewise distribution has a few drawbacks that perhaps limit its utility. First, several of the five parameters required may not be readily available or directly estimable. For example, the CMBL distribution uses F_2, the fraction of failures due to random failure. Instead of using F_2, λ_c may be more readily identifiable and available since most logistic/supply systems and experts use this constant failure rate for component lifetime estimations, even though this rate may inadvertently include portions of the infant mortality and wear-out failures. Second, the CMBL distribution allows no overlapping of the three sections of the lifecycle of the component, which theoretically is not correct although for the enterprise level models and PHM calculations, this may not be important. Third, the computation of internal parameters to model the probability distribution (PDF) requires numerical solving and is somewhat computationally intense. Fourth, the piecewise distribution is not theoretically continuous, although it can be considered approximately continuous for computational purposes. Finally, and most importantly, updating the parameters of the piecewise distribution through Bayesian methods raises difficult questions over how to proceed correctly, as demonstrated in the previous section.

4.2. Mathematical Formulation

4.2.1. Introduction

The Closed-form Full Life-Cycle (CFLC) probability density function is a function which represents failure probability density throughout the lifetime of a part. As the name indicates, the CFLC has a "closed-form" or "analytic" solution; it can be written in terms of known functions and constants. The CFLC combines the same three sections of the CMBL distribution, but has a slight change in its parameterization. The five intuitive and directly estimable input parameters are shown in Table 3. It is differentiable everywhere on the interval $[0,\infty)$, and integrates to unity on the same interval. The following sections provide the mathematical derivation of the CDF and PDF and describe the approach taken for parameter estimation.

Table 3. CFLC Distribution Parameter Summary

Parameter	Meaning	Valid Range
μ *	The mean of the wear-out portion of the distribution	$(0,\infty)$
σ	The standard deviation of the wear-out portion	$(0,\infty)$
λ	The failure rate due to random failures	$(0,\infty)$
Λ	The failure rate due to manufacturing or installation defects; a positive real	$(0,\infty)$
α	The proportion of items that have a manufacturing or installation defect that would eventually result in a failure if	$[0,1]$

	not precluded by another type of failure event.	

* μ should be at least 3σ, or the normal distribution is a poor representation of wear-out. In that case, a lognormal distribution may be more appropriate.

4.2.2. Mathematical Derivation

The derivation of the CFLC closed form PDF begins with several assumptions about the subject entity conveniently called a "part".

1. A known proportion of parts α will have one or more defect(s) that subject them to infant mortality related failures which may occur at any time, with a known constant failure rate Λ. In this case, α is equivalent to F_1 from the CMBL distribution.
2. Parts can fail randomly at any time, also with a known constant failure rate λ. In this case, λ is equivalent to λ_c in the CMBL distribution which is currently not an input parameter.
3. Parts can fail due to wear-out with a normally distributed lifetime with parameters μ and σ. These are the same input parameters as in the CMBL distribution.
4. All parts must eventually fail in exactly one of the three ways listed above.
5. The above failure modes are independent.

These assumptions warrant further discussion. The first assumption is believed to be valid, because no artificial constraint is put on the period of time in which an infant mortality failure might occur. Setting the failure rate high has the effect of decreasing the average time at which infant mortality occurs. Infant mortality is likely to occur early, but like a baby born with a heart valve problem that never is discovered until much later in life, the consequence of an inherent defect may be indefinitely delayed. The second assumption is relatively straightforward. Some failures are unpredictable, and memoryless with regard to time. They may be caused by interactions with other components or the external environment such as shocks to the system. The third assumption reflects the idea that most parts wear out eventually, even if no defect is present and no random failure ever occurs. Because a part can wear out in many ways, the central limit theorem would suggest that a normal distribution would be appropriate, although depending upon the system, a lognormal distribution may be appropriate. A good example of this is brake pads. Each application of the brake wears away a little bit of brake pad material, until in summation, these losses of material result in the complete elimination of the brake pad. The fourth assumption clarifies that only one of the three underlying failure events can occur. This is because the occurrence of one type of failure on a given part precludes the occurrence of any other type. The fifth assumption is the most important. For purposes of the derivation, the three types of failure are treated as independent events rather than as three possible outcomes of a single event. One way to think of this is that there are three different "clocks" running to keep track of the time at which a failure would occur if not precluded by a different type of failure.

For each type of failure event, we will define its probability distribution assuming independence from the other failure modes. Starting with the random failures, the PDF will be an exponential distribution, and the CDF is found by integration as shown below:

$$f_1(t) = \lambda e^{-\lambda t} \tag{21}$$

$$F_1(t) = \int_0^t f_1(v)dv = 1 - e^{-\lambda t} \tag{22}$$

Next are the PDF and CDF for the normal distribution. "Erf" is the well known error function for which numerical implementations are widely available.

$$f_2(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}} \tag{23}$$

$$F_2(t) = \int_0^t f_2(v)dv = \frac{1}{2}\left(Erf\left(\frac{t-\mu}{\sqrt{2}\sigma}\right) + Erf\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right) \tag{24}$$

Lastly, we present the PDF and CDF of the infant mortality distribution. The PDF here is exponential, as in the case of random failure, but the additional parameter α has been multiplied with the usual exponential PDF. Note that in this case, the CDF tends to α as $t \to \infty$. This is as expected, since if infant mortality were the only failure mode, only the percentage of parts that begin life in a defective state would ever fail.

$$f_3(t) = \alpha \Lambda e^{-\Lambda t} \tag{25}$$

$$F_3(t) = \int_0^t f_3(v)dv = \alpha(1 - e^{-\lambda t}) \tag{26}$$

Now that we've defined the three underlying distributions that will go into the full life-cycle distribution, we will combine them using the addition law of probability. The addition law of probability tells us that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{27}$$

or with three variables,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \tag{28}$$

Below is a representation of the associated Venn diagram:

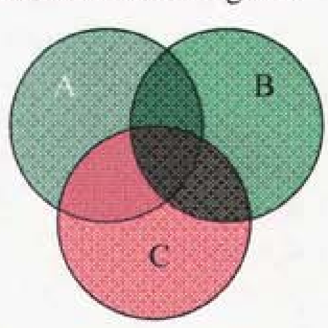


Figure 9. Venn Diagram

Now we will use "F" to represent the CDF of the desired distribution, and "f" to represent the associated PDF. F(t) should represent the probability that any type of failure will occur in the time interval $(0,t)$. Thus $F(t)$ should be the probability of an F_1, F_2 or F_3 type failure. Applying the addition law of probability and using the assumption of independence yields:

$$F(t) = F_1(t) + F_2(t) + F_3(t) - F_1(t)F_2(t) - F_1(t)F_3(t) - F_2(t)F_3(t) + F_1(t)F_2(t)F_3(t) \tag{29}$$

For a function to meet the definition of a CDF, it must monotonically increase from 0 to 1 as the independent variable, t, goes from 0 to infinity. Equivalently, the resulting PDF must be non-negative everywhere and integrate to unity on the interval from 0 to infinity. It is of interest here that defining the CDF in this way (using the addition law of probability and independence) guarantees that the resulting CDF and PDF will represent a valid distribution.

$$
\begin{aligned}
&F(\infty) = F_1(\infty) + F_2(\infty) + F_3(\infty) - F_1(\infty)F_2(\infty) - F_1(\infty)F_3(\infty) - F_2(\infty)F_3(\infty) + F_1(\infty)F_2(\infty)F_3(\infty) \\
&F(\infty) = 1 + 1 + \alpha - 1\cdot 1 - 1\cdot\alpha - 1\cdot\alpha + 1\cdot 1\cdot\alpha \\
&F(\infty) = 1 + 1 + \alpha - 1 - \alpha - \alpha + \alpha \\
&F(\infty) = 1
\end{aligned}
\tag{30}
$$

As a matter of fact, it can be shown that any number of PDFs could be combined in this manner and if at least one of them integrates to unity, so will the resulting PDF.

To show that F is a monotonically increasing function on the interval 0 to infinity, we need only show that f is non-negative. For any arbitrary set $\{f_1, f_2, f_3,\}$ of PDFs and their associated set of CDFs, $\{F_1, F_2, F_3\}$, $f(t)$ expands to:

$$
\begin{aligned}
f(t) = {} & f_1(t) + f_2(t) + f_3(t) - f_1(t)F_2(t) - F_1(t)f_2(t) - f_1(t)F_3(t) - F_1(t)f_3(t) \\
& - f_2(t)F_3(t) - F_2(t)f_3(t) + f_1(t)F_2(t)F_3(t) + F_1(t)f_2(t)F_3(t) + F_1(t)F_2(t)f_3(t)
\end{aligned}
\tag{31}
$$

By factoring out the f_1, f_2, and f_3, terms, we can see that $f(t)$ is the sum of three non-negative terms, and is therefore non-negative.

$$
\begin{aligned}
f(t) = {} & \{(1 - F_2(t))(1 - F_3(t))\}f_1(t) + \{(1 - F_1(t))(1 - F_3(t))\}f_2(t) + \\
& \{(1 - F_1(t))(1 - F_2(t))\}f_3(t)
\end{aligned}
\tag{32}
$$

Therefore, we can state that any $f(t)$ derived in this manner will be a valid PDF, so long as the inputs are non-negative, and at least one of them integrates to unity.

Substitution of the functions $\{F_1, F_2, F_3\}$ chosen for our application and simplification yields the CDF

$$
F(t) = 1 + \frac{1}{2}e^{-t(\lambda+\Lambda)}\left(e^{t\Lambda}(-1+\alpha) - \alpha\right)\left(Erfc\left(\frac{t-\mu}{\sqrt{2}\sigma}\right) + Erfc\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right) \tag{33}
$$

Note that "Erfc" is the complementary error function.

Differentiation yields the CFLC PDF

$$
f(t) = F'(t) \tag{34}
$$

$$
f(t) = \frac{1}{2}e^{-t(\lambda+\Lambda)}\left(\frac{e^{\frac{-(t-\mu)^2}{2\sigma^2}}\sqrt{2/\pi}\left(-e^{t\Lambda}(-1+\alpha)+\alpha\right)}{\sigma} + \left(e^{t\Lambda}(-1+\alpha)\lambda - \alpha(\lambda+\Lambda)\right)\left(-2 + Erf\left(\frac{t-\mu}{\sqrt{2}\sigma}\right) + Erf\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right)\right) \tag{35}
$$

The above equations for the PDF and CDF have been extensively simplified using Mathematica.

4.2.3. Example Plots

A plot of the PDF of the resulting distribution with some arbitrarily chosen numerical parameters is shown in Figure 10. The term nf(t) refers to an instance of f(t) with numerically defined input parameters. Notice that the plot has the same features found in the CMBL distribution.

$$nf(t) = f(t)|(\mu = 10, \sigma = 1, \lambda = 1/10, \Lambda = 10, \alpha = 1/20) \tag{36}$$

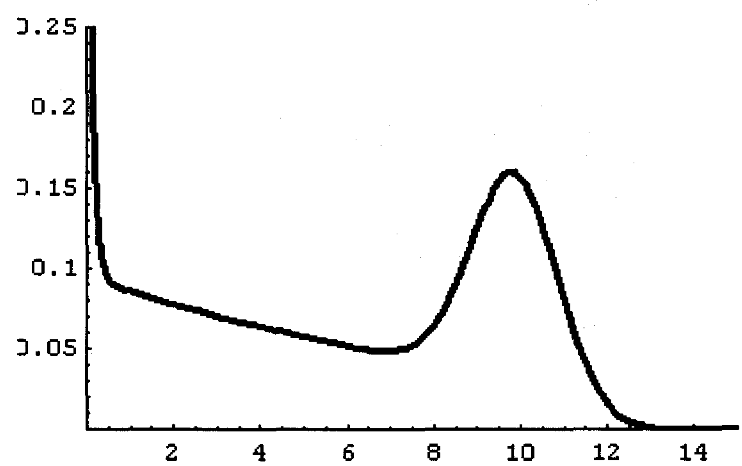


Figure 10. Example CFLC PDF

The plot of the CFLC CDF is shown in Figure 11.

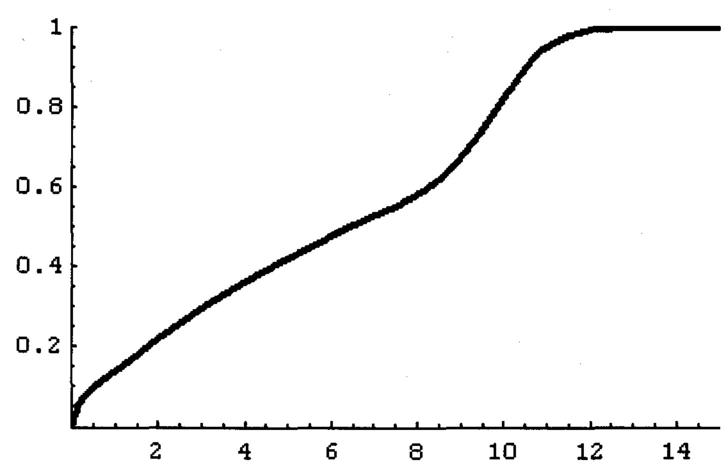


Figure 11. Example CFLC CDF

The hazard function defined by $h(t) = \dfrac{f(t)}{1 - F(t)}$ is shown in Figure 12 and yields a shape similar to the familiar "bathtub" curve.

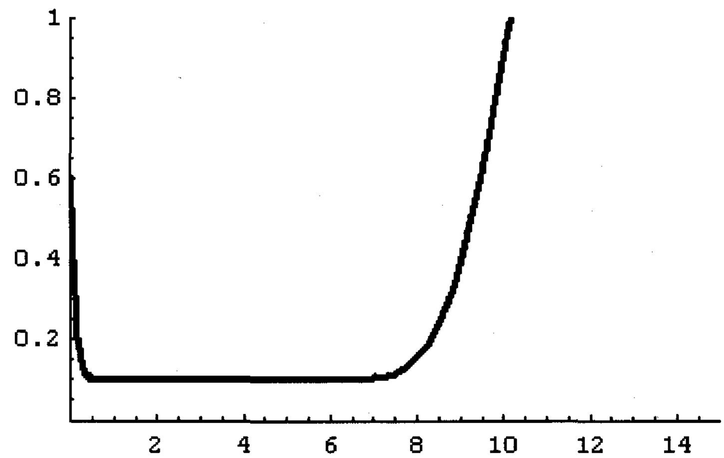


Figure 12. Example CFLC Hazard Function

It is apparent from the above plots that the CFLC distribution has the same important qualities observed in the CMBL distribution – namely that the three failure intervals known to occur in real data are easily observed. It is also apparent that the CFLC distribution has a hazard function similar to the "bathtub" curve, which is often modeled as a piecewise set of functions as shown below.

$$y(t) = \begin{cases} c_0 - c_1 t + \lambda, & 0 \le t \le c_0/c_1 \\ \lambda, & c_0/c_1 < t \le t_0 \\ c_2(t - t_0) + \lambda, & t_0 < t \end{cases} \tag{37}$$

One obvious advantage of the distribution presented here over the piecewise methods such as the bathtub curve or the CMBL distribution is that if the three input PDFs are differentiable everywhere, the resulting distribution will be differentiable everywhere also, as a consequence of Equation (32).

4.2.4. Parameter Estimation

Typically, one needs to know, given the five input parameters, what is the mean and standard deviation for a distribution. This knowledge allows side by side comparison of different distributions. The mean of the closed-form full life-cycle distribution is given by:

$$mean = E[T] = \int_0^\infty t\, f(t)dt \tag{38}$$

where T is the random variable of time to failure. After much algebraic manipulation, this reduces to:

$$mean = \frac{2(\lambda + \Lambda - \alpha\Lambda) + e^{\frac{1}{2}\lambda(-2\mu + \lambda\sigma^2)}(-1+\alpha)(\lambda+\Lambda)Erfc\left(\frac{-\mu + \lambda\sigma^2}{\sqrt{2}\sigma}\right) - e^{\frac{1}{2}(\lambda+\Lambda)(-2\mu + (\lambda+\Lambda)\sigma^2)}(\alpha\lambda)Erfc\left(\frac{-\mu + (\lambda+\Lambda)\sigma^2}{\sqrt{2}\sigma}\right)}{2\lambda(\lambda+\Lambda)} \tag{39}$$

The standard deviation can be computed as follows:

$$std.dev. = \sqrt{E[(T - E[T])^2]} = \int_0^{\infty} (t - mean)^2 f(t)\,dt \qquad (40)$$

Unfortunately, the problem of producing a closed form solution for the standard deviation seems to be an intractable problem. However, when the values of the input parameters are known, the standard deviation is easily calculated using numerical integration. In addition, the mean and standard deviation for this type of distribution may not have any real-world relevance since the distribution is multimodal.

Symbolically deriving estimators for the parameters has also proven to be very difficult and has not been accomplished at the time of publishing. However, using the method of Maximum Likelihood Estimation with numerical optimization has been demonstrated to yield very suitable parameter estimates. Results were obtained by using Equation (35) to generate a randomly drawn list of 100 data points. The method of maximum likelihood estimation was implemented using numerical optimization yielding the parameter estimates shown in Table 4. A graphical comparison of the Original PDF (shown in green) and the Estimated PDF (shown in blue) using 100 Monte Carlo generated data points is presented in Figure 13.

Table 4. Parameter Estimation Using 100 Monte Carlo Generated Data Points

Parameter	True Value	Estimated Value	% Error
μ	10.00	9.77	-2.3%
σ	1.00	1.35	35%
λ	0.1000	.0778	-22%
Λ	10.00	8.32	-17%
α	.050	.147	190%

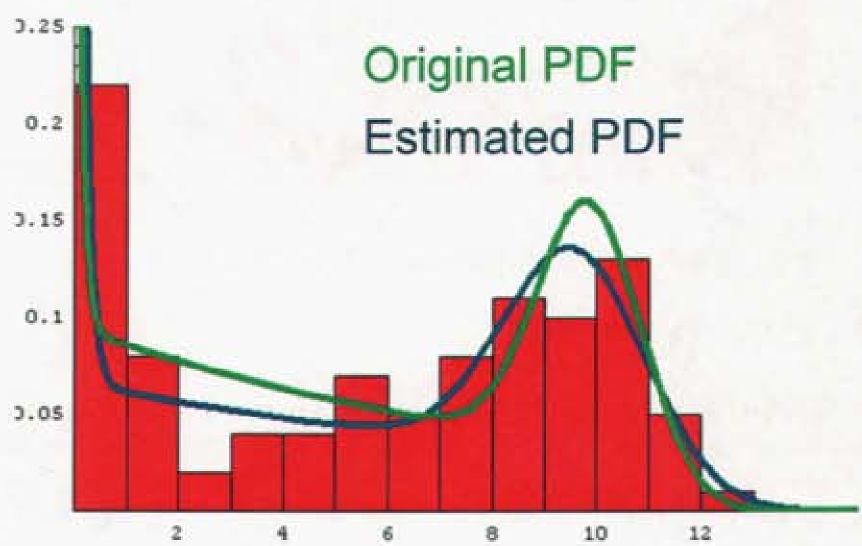


Figure 13. Comparison of 100 Data Points

When the number of Monte Carlo generated data points is increased to 1000, the error in the estimates is reduced considerably. Results were again obtained by using Equation (35) to generate a randomly drawn list of 1000 data points. The method of maximum likelihood estimation was implemented using numerical optimization yielding the parameter estimates shown in Table 5. A graphical comparison of the Original PDF (shown in green) and the Estimated PDF (shown in blue) using 1000 Monte Carlo generated data points is presented in Figure 14. The parameters for the infant mortality prove to be the most difficult to estimate because a small portion of the generated data falls into that domain. The fact that a large number of data points are needed in order to establish a good approximation for the distribution, combined with the reality that such extensive failure data is seldom available, makes a strong argument that a Bayesian approach is needed for estimating the parameters of the distribution.

Table 5. Parameter Estimation Using 1000 Monte Carlo Generated Data Points

Parameter	True Value	Estimated Value	% Error
μ	10.00	9.90	-1%
σ	1.00	1.06	6%
λ	0.1000	.105	5%
Λ	10.00	14.91	49%
α	.050	.054	8%

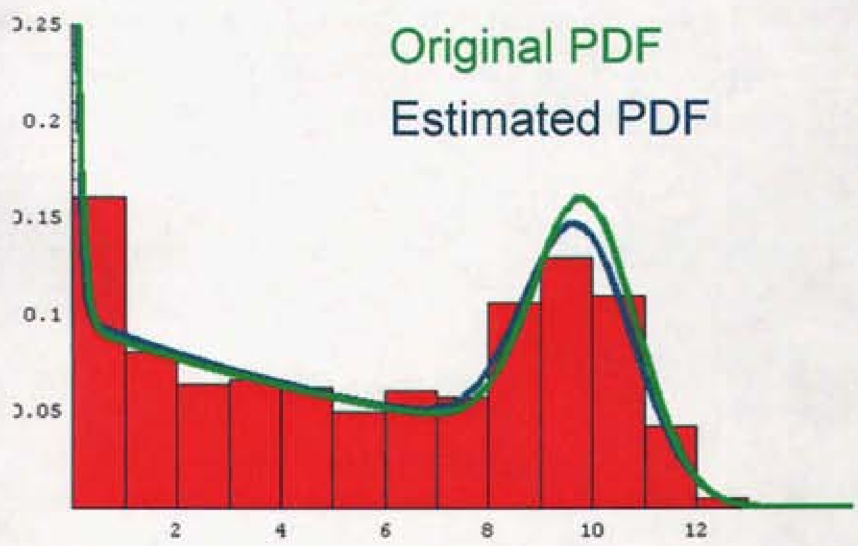


Figure 14. Comparison of 1000 Data Pts

4.3. CFLC Distribution Summary

The Closed-form Full Life Cycle distribution was developed using three distributions treated as modeling independent events. The CFLC CDF, Equation (33), of the resulting distribution was derived using the addition law of probability on the original CDFs. The resulting CFLC PDF distribution, Equation (35), has the properties needed to model the full life cycle of a typical part, and has advantages over piecewise methods that attempt to do the same. Because the distribution has a closed form solution, it can take advantage of several different methods for updating it's parameters based on new data from either field failures or sensors. However, additional theoretical validation of the CFLC distribution is warranted prior to implementation, since this was an initial attempt at a closed-form solution.

5. CONCLUSIONS

The CMBL distribution provides an application friendly method for characterizing a component's failure or lifecycle distribution. This paper characterized the distribution and explored methods for updating the CMBL distribution as new data become available. The initial results obtained in applying the Bayesian sequential updating methodology to the CMBL distribution shows some promise. However, additional research is needed for the updating process involved with the random portion of the curve.

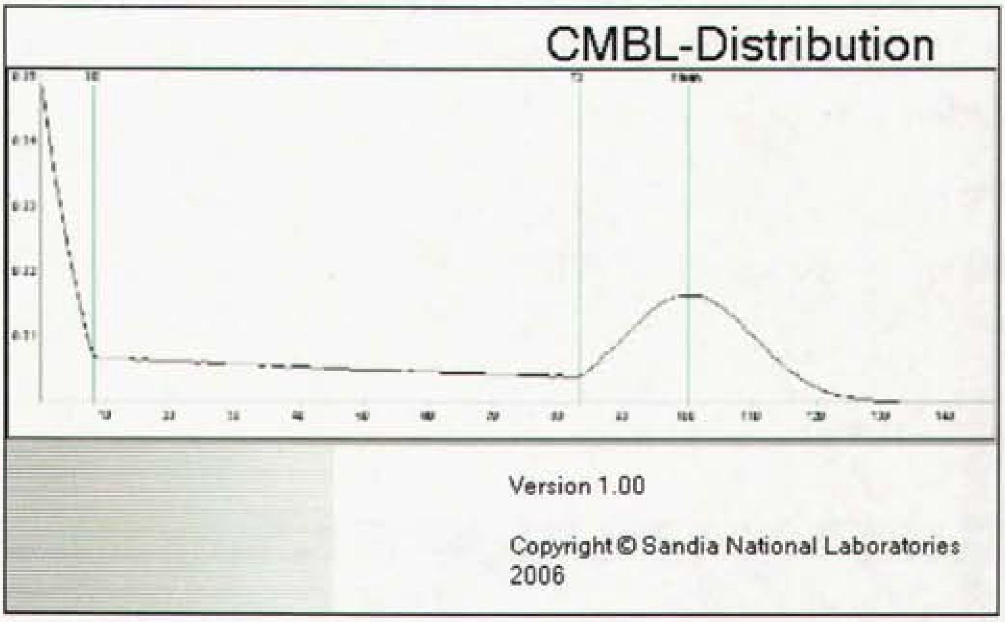
The development of the CFLC distribution shows great promise. Being a closed-form solution, updating methods other than that attempted with the CMBL distribution may provide a more computer efficient result. Further research is warranted to both provide theoretical validation and determine which updating method would be more appropriate for enterprise level logistics and PHM modeling.

These two Sandia National Laboratories' methods developed in this effort for updating the CMBL distribution and other TTF distributions, should be valuable in enhancing maintenance planning and real-time situational awareness processes. Following additional exploration and validation, these methods, used in enterprise level and prognostics and health management PHM modeling, should more accurately help provide timely feedback on the current status of equipment; provide tactical assessment of the readiness of equipment for the next campaign; identify parts, services, etc. that are likely to be required during the next campaign; provide a realistic basis for scheduling and optimizing equipment maintenance schedules; and help ensure that the useful life of expensive components is maximized while reducing the incidence of unplanned maintenance.

6. REFERENCES

1. L. C. Wolstenholme, Reliability Modelling, A Statistical Approach, Chapman & Hall/CRC, New York, 1999.
2. H. Caruso, Too Much Time in the Bathtub (Curve) – "An Aging Aircraft Paradigm that Doesn't Hold Water." Joint Council on Aging Aircraft, Aging Aircraft Conference Technical Paper, 2005.
3. D. J. Wilkins, "The Bathtub Curve and Product Failure Behavior Part Two – Normal Life and Wear-Out," Reliability Hot Wire, Issue 22, December 2002.
4. J. M. Booker, et al., "PREDICT: A New Approach To Product Development and Lifetime Assessment Using Information Integration Technology," Submitted to Handbook of Statistics: Statistics in Industry, September, 2000.
5. M. Xie, C. C. Lai, "Reliability Analysis Using an Additive Weibull Model with Bathtub-Shaped Failure Rate Function," Reliability Engineering and System Safety, Vol 52, (1995), pp 87-93.
6. Future Combat Systems Integrated Support Team Accomplishments. Sandia Lab News, Vol 57, No. 25, 9 December, 2005, 8.
7. System of Systems Analysis Toolset (SoSAT) User's Manual. Sandia National Laboratories, Division 6642, May 2006.
8. Swiler, Laura P., James E. Campbell, et al. Algorithm Development for Prognostics and Health Management (PHM), SAND Report, SAND2003-3820, October, 2003.
9. Campbell, James E. Calculation of the Inverse Wear-out Distribution. Unpublished paper, Sandia National Laboratories, June, 1998.
10. H. F. Martz, R. A. Waller, *Bayesian Reliability Analysis,* John Wiley & Sons, New York, 1982.

APPENDIX A: CMBL-DISTRIBUTION GRAPHER



A.1. Overview

The CMBL-Distribution Grapher displays a graph based on the input parameters for a CMBL distribution and allows the user to explore the valid range in three dimensions of the CMBL distribution. The following subsections provide a detailed look at the input, output, and capabilities of the CMBL-Distribution Grapher.

A.2. Plotting a graph

A.2.1. Step 1

The first step in plotting a graph is to enter the five parameters for the CMBL distribution:

1. Mean
2. Standard Deviation
3. Burn in Fraction
4. Burn in Duration
5. Random Fraction

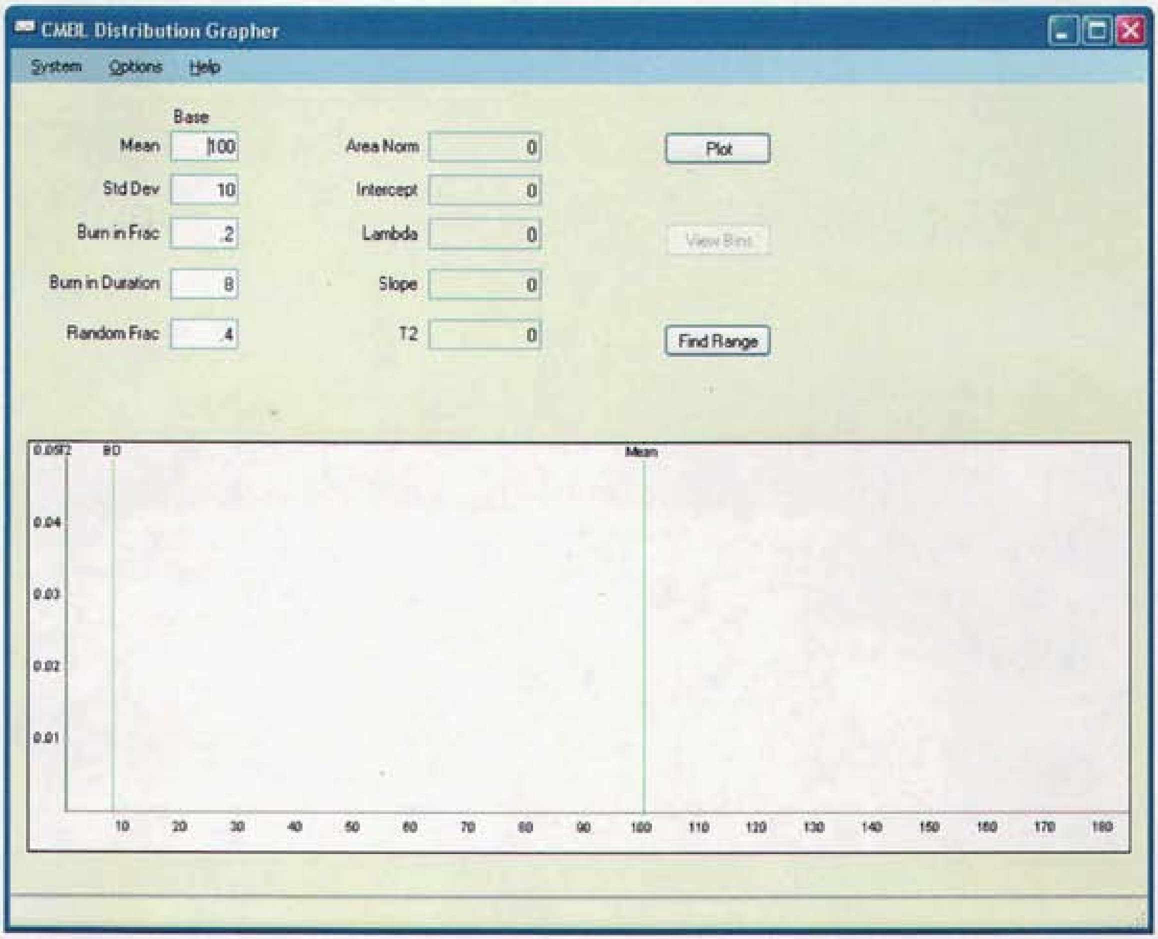


Figure 15. Completed Parameter Input

A.2.2. Step 2

After all parameters have been entered, clicking the Plot button will display the CMBL distribution in the graph window as shown in Figure 16. In the event of an illegal parameter, the graph window is cleared and an error message is displayed.

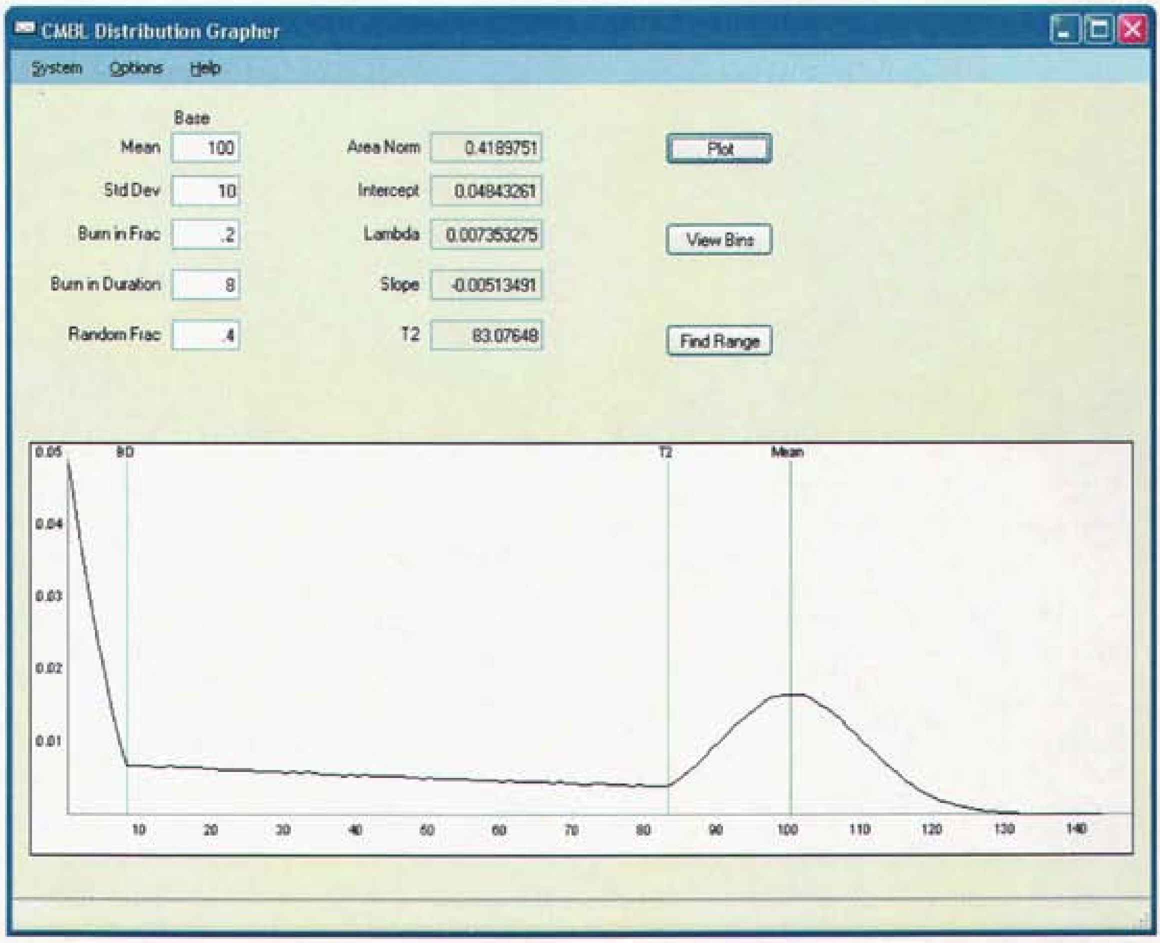


Figure 16. Plotted CMBL distribution

A.2.3. Step 3

There are three options to control the speed and resolution of the graph (all times are specific to a 3Ghz Pentium) as shown in Table 6 and Figure 17.

Table 6. Graph Resolution

Plot	Samples	Time (seconds)
Fast	1,000	< 1
Normal	100,000	1
Fine	250,000	2
Very Fine	1,000,000	3

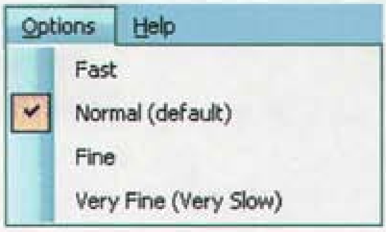


Figure 17. Options Menu

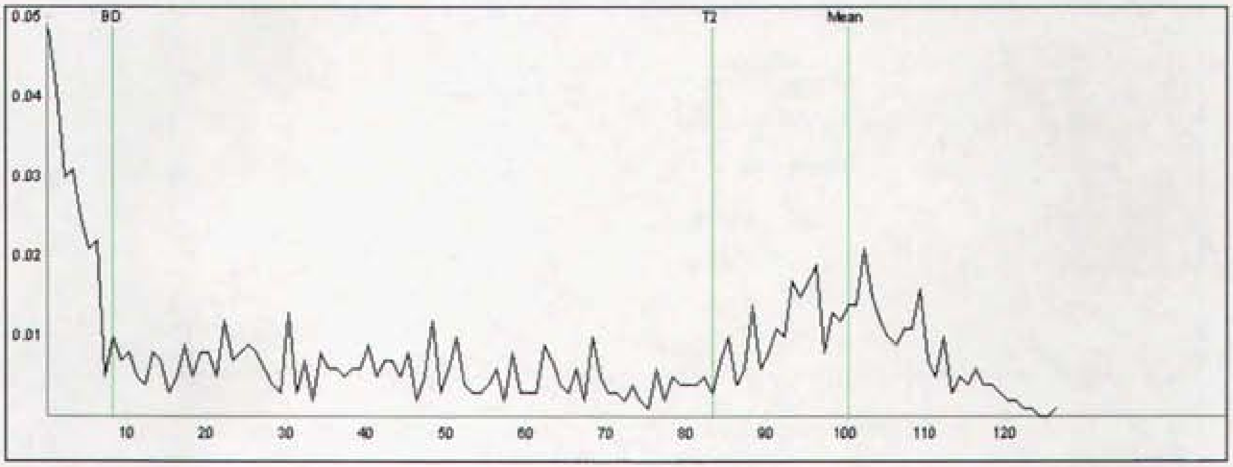


Figure 18. Fast Option

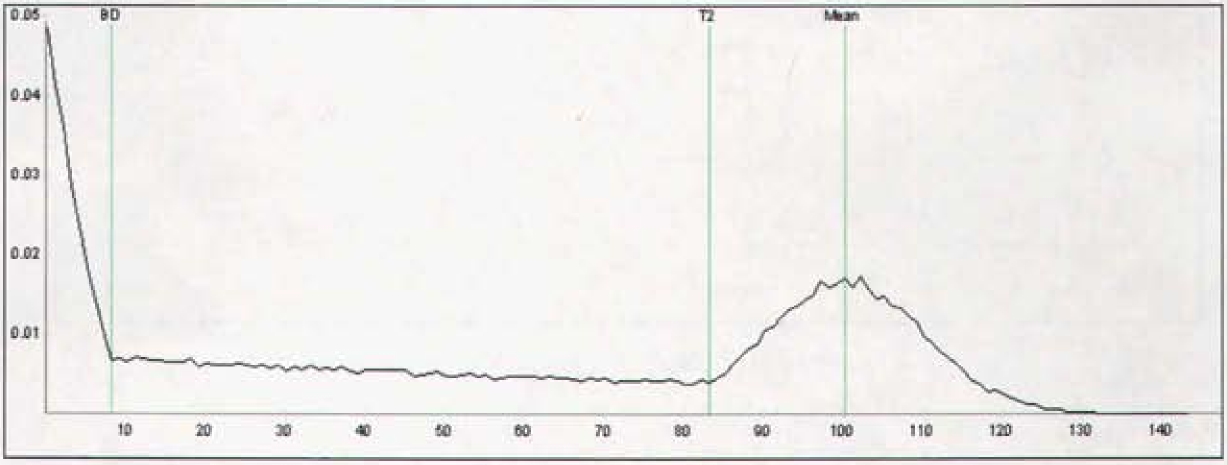


Figure 19. Normal Option

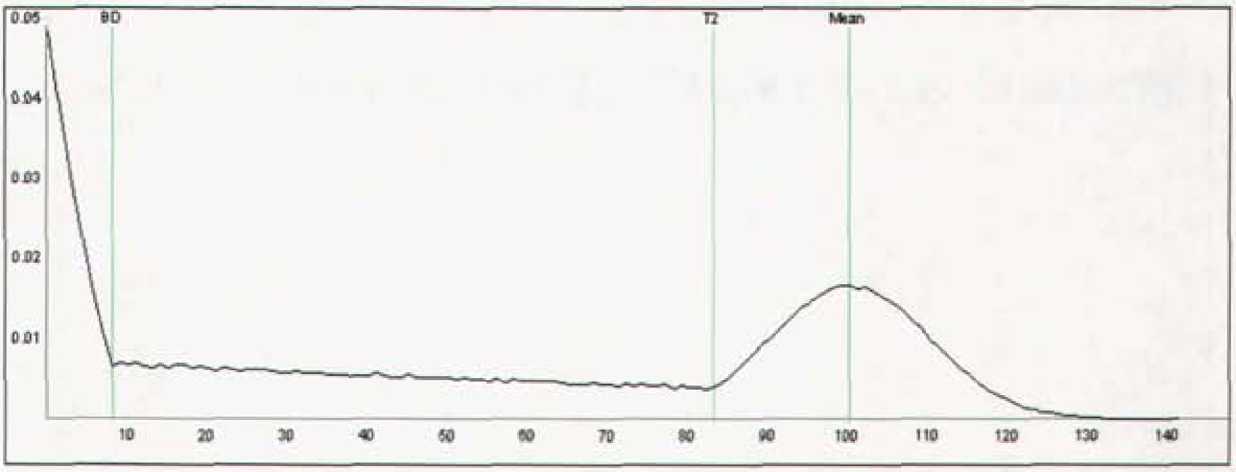


Figure 20. Fine Option

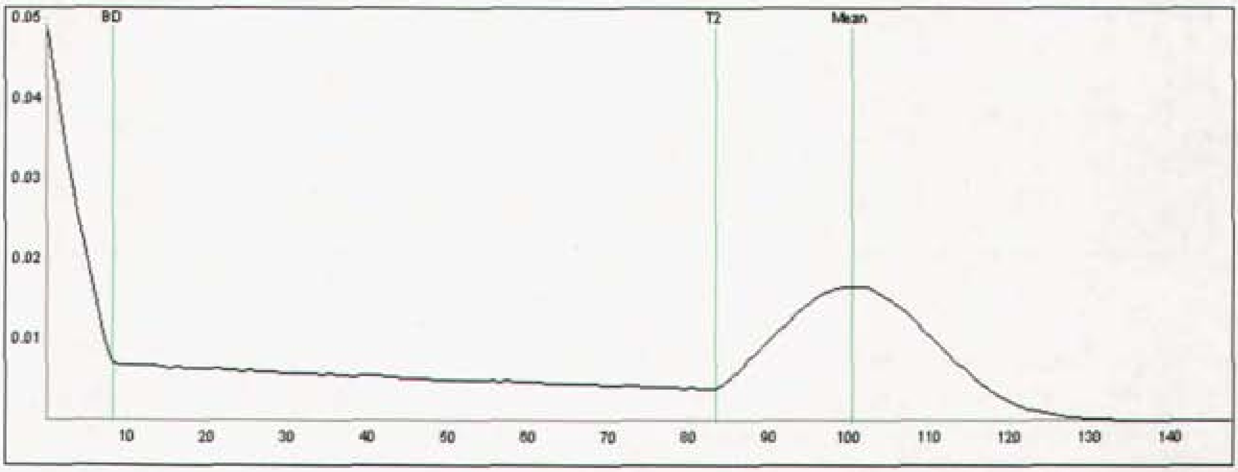


Figure 21. Very Fine Option

A.3. Viewing and Exporting the Bins

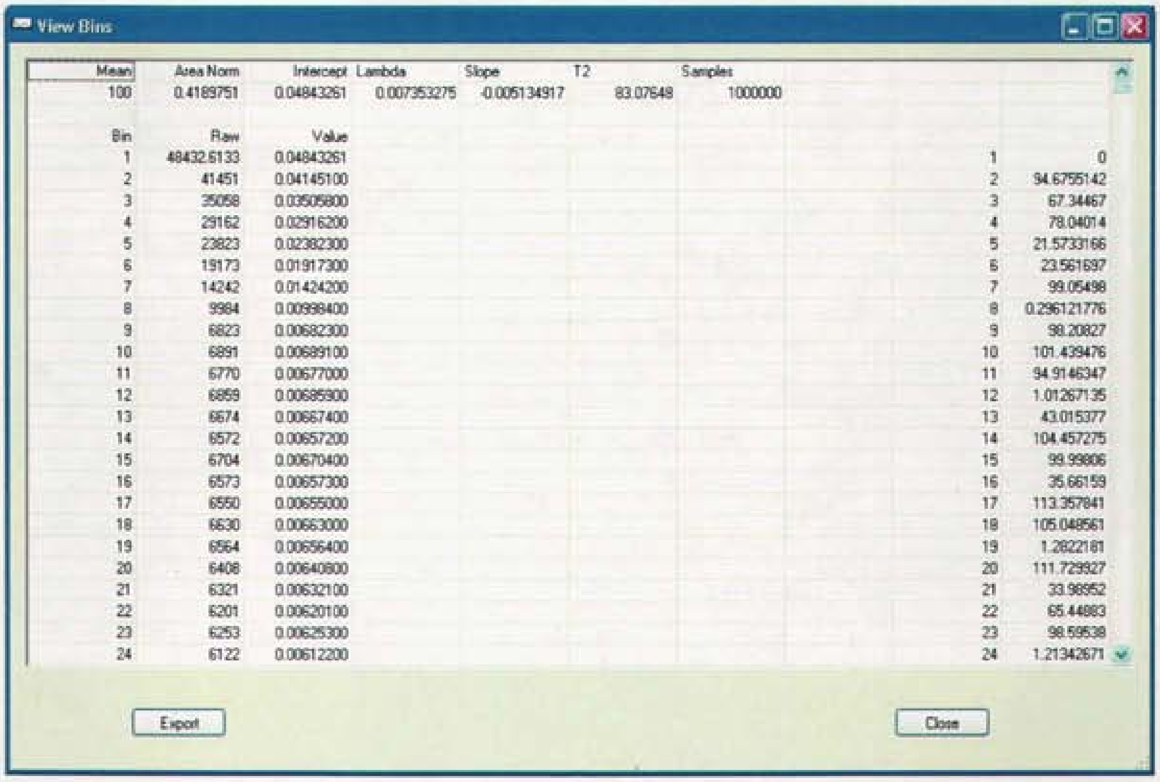


Figure 22. View Bins

The "View Bins" button will be enabled after successfully plotting a CMBL distribution. Clicking the "View Bins" button from the main form will display the data bins used to generate the plot. Data on the View Bins form is read-only. Clicking the "Export" button creates a comma separated file formatted for importing into and Excel spreadsheet. Rows 1 and 2 contain information relating directly to the distribution as shown in Table 7.

Table 7. Distribution Information

Column	Value
1	Mean
2	Area Norm
3	Intercept
4	Lambda
5	Slope
6	T2
7	Samples

Rows 3 and higher contain bin and raw data as shown in Table 8.

Table 8. Bin and Raw Data Information

Column	Value

1	Bin
2	Raw Total
3	Value
9	Raw Index
10	Raw Value

A.4. Validating Ranges

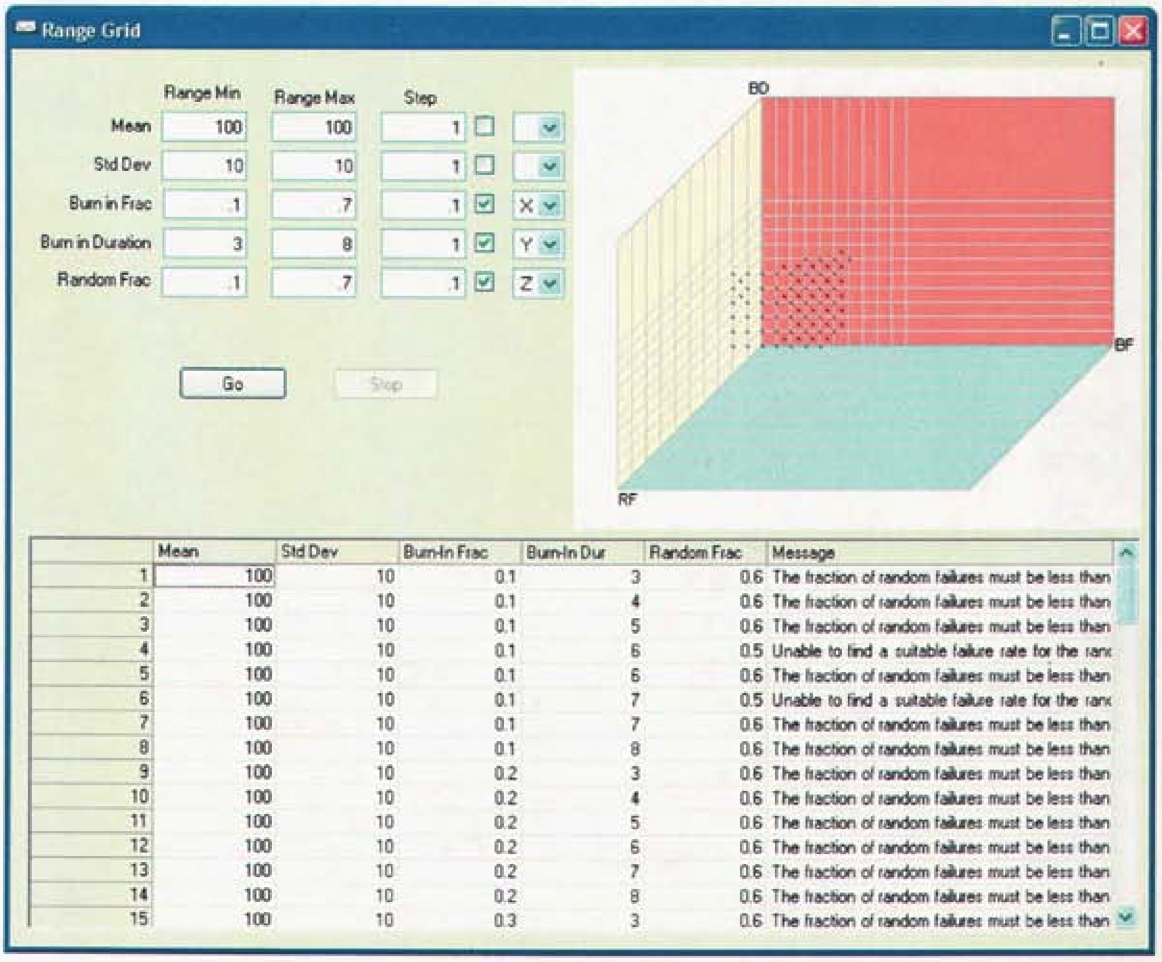


Figure 23. Range Form

The purpose of the Range form is to find the boundaries of the CMBL-Distribution by modifying the input range and plotting the values that trigger an error in the distribution. The errors will be plotted in the xyz chart and will be identified in the grid. All five of the input ranges may be used for a run but only three may be plotted at any one time. The ranges to plot are selected by adding a checkmark after the range and identifying which axis to use for that range.

Inputs:

1. Mean
2. Standard Deviation
3. Burn-In Fraction

4. Burn-In Duration
5. Random Fraction

Values:

1. Range Min – Minimum value for this range
2. Range Max – Maximum value for this range
3. Step – Increment value for this range. The range will not increment above the maximum value for the range.

Outputs:

1. Mean
2. Standard Deviation
3. Burn-In Fraction
4. Burn-In Duration
5. Random Fraction
6. Message

A.5. Requirements

A.5.1. Minimum

Windows XP Home or later
Pentium Class 1 GHz processor
128 MB RAM
15 MB free hard drive space (40 MB if installing DotNet 2.0 Framework)
DotNet Framework 2.0

A.5.2. Suggested

Windows XP Pro or later
Pentium Class 3 GHz processor
256 MB RAM
15 MB free hard drive space (40 MB if installing DotNet 2.0 Framework)
DotNet Framework 2.0

DISTRIBUTION

MS 1005	Russell. D. Skocypec (06340)
MS 1005	Dennis Engi (06340)
MS 1011	Robert M. Cranwell (6342)
MS1011	Daniel Briand (6342)
MS1011	Kelly S. Lowder (6342)
MS1011	Donald N. Shirah (6342)
MS9018	Central Technical Files (8944) (2)
MS0899	Technical Library (4536) (2)
MS0123	D. Chavez, LDRD Office (1011)